Theorem 3 [Updated version 03 · 15 · 2012

Consider the functional $g : X \to \mathbb{R}$ on a normed linear space X. Let

$$\Omega = \{ x \in X : g_i(x) = 0, \ i = 1, 2, \cdots, n \}$$

be a constraint set defined by the functionals $g_i : X \to \mathbb{R}$, $i = 1, 2, \cdots, n$.

Assume g and g_i, $i = 1, 2, \ldots, n$ are Fréchet differentiable.

Suppose x_0 is an extremum of g subject to the constraints $g_i(x) = 0$ $i = 1, 2, \cdots, n$, and x_0 is a regular point of Ω. Then,

$$\bigcap_{i=1}^{n} \operatorname{Ker}(Dg_i(x_0)) \subset \operatorname{Ker} Dg(x_0) .$$

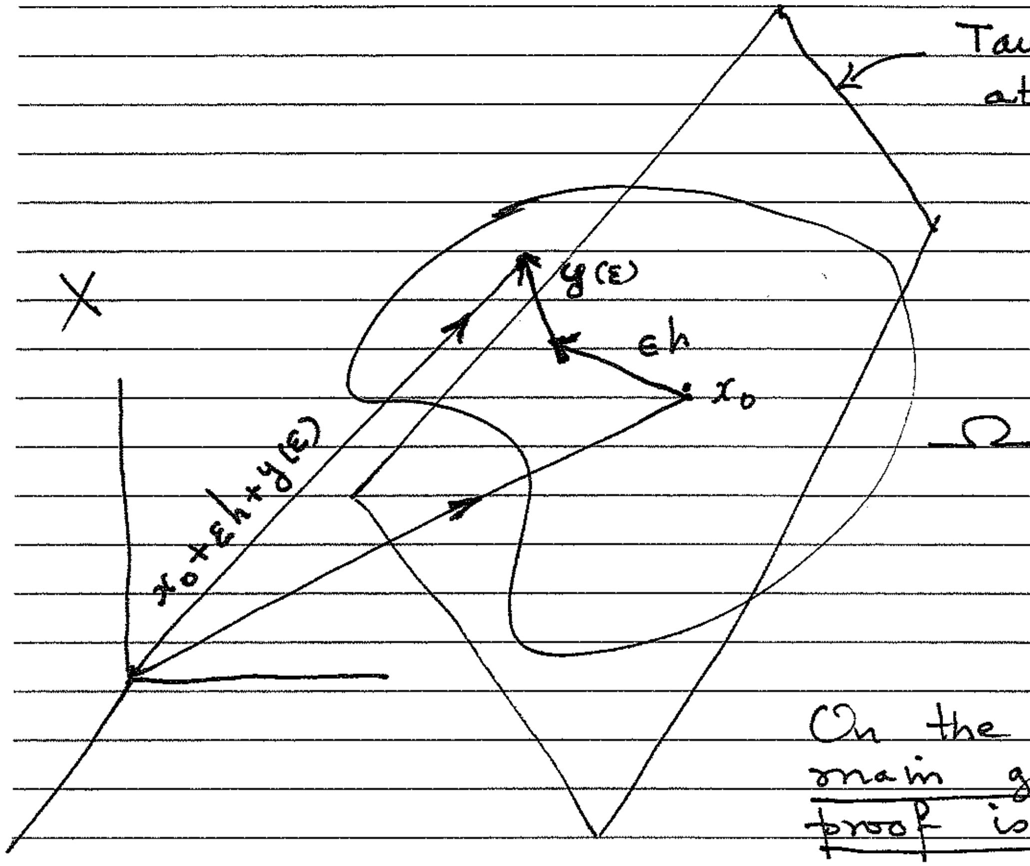
Proof : Let $h \in \bigcap_{i=1}^{n} \operatorname{Ker}(Dg_i(x_0))$.

By Theorem 1 and associated remark (Lect 5(b)) there exist n linearly independent vectors $y_1, y_2, \cdots, y_n \in X$ such that

$$M = [Dg_i(x_0) y_j] = \mathbb{1}_n \text{ the } n \times n$$

Refer to the figure ~~on the~~ next page depicting the constraint set as a manifold or hypersurface Ω. The plane containing x_0 is also the tangent plane to ~~Ω~~ at x_0. By hypothesis, h and hence ϵh for any scalar ϵ belongs to this tangent plane. < Think about proving this >

Tangent plane to Ω at x_0 denoted $T_{x_0}\Omega$



By hypothesis

$h \in$ tangent plane

But

$x_0 + \varepsilon h \notin \Omega$.

On the other hand the <u>main goal of this proof is to show</u> that

there exists a suitable $y(\varepsilon)$ such that

$$x_0 + \epsilon h + y(\varepsilon) \in \Omega$$

provided ϵ is small enough. For this, consider the system of equations

$$g_i \left(x_0 + \epsilon h + \sum_{i=1}^{n} \varphi_i y_i \right) = 0$$

$i = 1, 2, \ldots, n$.

where x_0, h, y_i, $i=1,2,\ldots,n$ are fixed as above, but ϵ and φ_i $i=1,2,\ldots,n$ are $(n+1)$ real valued unknowns.

Define $\varphi = \begin{pmatrix} \varphi_1 \\ \varphi_2 \\ \vdots \\ \varphi_n \end{pmatrix}$ and $\tilde{g} = \begin{pmatrix} g_1 \\ g_2 \\ \vdots \\ g_n \end{pmatrix}$.

The above equations take the form

$$\tilde{g} : \mathbb{R} \times \mathbb{R}^n \to \mathbb{R}^n$$

$$(\epsilon, \varphi) \longmapsto \tilde{g}(\epsilon, \varphi) = \begin{pmatrix} g_1(x_0 + \epsilon h + \sum_i \varphi_i y_i) \\ \vdots \\ g_n(x_0 + \epsilon h + \sum_i \varphi_i y_i) \end{pmatrix}$$

$$= 0 \in \mathbb{R}^n$$

Clearly $\tilde{g}(0,0) = 0$ (since $x_0 \in \Omega$)

Observe that the partial Fréchet derivative

$$D_2 \tilde{g}(0,0) = [D g_i(x_0) y_j] = \mathbb{1}_n$$

invertible. Hence the implicit function theorem applies at $\epsilon = 0$, $\varphi = 0$, and there is an open interval

$$U = (-\varepsilon_0, \varepsilon_0)$$

containing 0, such that $\forall\, \epsilon$ in U, there is a unique vector-valued function $\epsilon \mapsto \varphi(\epsilon)$ such that

$$\tilde{g}(\varepsilon, \varphi(\varepsilon)) = 0$$

Then, for $y(\varepsilon) = \sum_{i=1}^{n} \varphi_i(\varepsilon) y_i$,

$$g_i(x_0 + \varepsilon h + y(\varepsilon)) = 0 \qquad i = 1, 2, \dots, n.$$

(small 'oh')

We let any expression $o(\varepsilon)$ to mean that

$$\lim_{\varepsilon \to 0} \frac{o(\varepsilon)}{\varepsilon} = 0.$$

Also, clearly $\lim\limits_{\varepsilon \to 0} \| \varepsilon h + y(\varepsilon) \| = 0$ (recall $y(0)=0$)

By definition of Fréchet derivative

$$\lim_{\varepsilon \to 0} \frac{\| g_i(x_0 + \varepsilon h + y(\varepsilon)) - g_i(x_0) - Dg_i(x_0)(\varepsilon h + y(\varepsilon)) \|}{\| \varepsilon h + y(\varepsilon) \|}$$

$$= 0$$

Equivalently, for $i = 1, 2, \cdots, n$,

$$g_i(x_0 + \varepsilon h + y(\varepsilon)) - g_i(x_0)$$

$$= Dg_i(x_0)(\varepsilon h + y(\varepsilon)) + o(\|\varepsilon h + y(\varepsilon)\|)$$

$$-\varepsilon_0 < \varepsilon < \varepsilon_0$$

But the left hand side $= 0$ by construction.

Also $Dg_i(x_0)h = 0$. Hence, using

$y(\varepsilon) = \sum\limits_{j=1}^{n} \varphi_j(\varepsilon) y_j$, we note,

$$0 = \left[D g_i (x_0) y_j \right] \varphi(\varepsilon) + o(\varepsilon) + o(\|y(\varepsilon)\|)$$

$$= \varphi(\varepsilon) + o(\varepsilon) + o(\|y(\varepsilon)\|)$$

We need the following calculations.

dot product

$$y(\varepsilon) = \sum_{i=1}^{n} \varphi_i(\varepsilon) y_i = \sum_{i=1}^{n} (e_i \cdot \varphi(\varepsilon)) y_i$$

where $e_i = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} \leftarrow i^{th}$ place

(i) $$\Rightarrow \|y(\varepsilon)\| \le \sum_{i=1}^{n} |e_i \cdot \varphi(\varepsilon)| \, \|y_i\|$$

$$\le \sum \|e_i\|_2 \, \|\varphi(\varepsilon)\|_2 \, \|y_i\|$$

(Cauchy-Schwartz)

$$= \left(\sum_{i=1}^{n} \|y_i\|_{\cancel{2}} \right) \|\varphi(\varepsilon)\|_2$$

$$\le \frac{1}{d_1} \|\varphi(\varepsilon)\|$$

← any norm, and suitable d_1 (by equivalence of all norms in $\mathbb{R}^n$)

$$\Leftrightarrow \quad d_1 \|y(\varepsilon)\| \le \|\varphi(\varepsilon)\| \quad ✓$$

(ii) $L : \mathbb{R}^n \longrightarrow [y_1, y_2, \dots, y_n]$ = closed linear span of the linearly independent vectors inside square brackets (of dimension n)

$$\varphi \longmapsto \sum \varphi_i y_i$$

is invertible (one-to-one otherwise y_i $i=1,2,\dots,n$ is not a linearly independent set)

$$\varphi(\varepsilon) = L^{-1} L \varphi(\varepsilon) = L^{-1} y(\varepsilon)$$

$$\Rightarrow \quad \|\varphi(\varepsilon)\| = \|L^{-1} y(\varepsilon)\|$$

$$\leq \|L^{-1}\| \cdot \|y(\varepsilon)\|$$

$$= d_2 \|y(\varepsilon)\| \quad \checkmark$$

Combining (i) and (ii)

(iii) $$d_1 \|y(\varepsilon)\| \leq \|\varphi(\varepsilon)\| \leq d_2 \|y(\varepsilon)$$

But we have shown that,

$$0 = \varphi(\varepsilon) + o(\varepsilon) + o(\|y(\varepsilon)\|)$$

Thus $\|y(\varepsilon)\| = o(\varepsilon)$, using ~~(i) and ii)~~ (iii)

for $-\varepsilon_0 < \varepsilon < \varepsilon_0$.

Since $x_0 + \varepsilon h + y(\varepsilon) \in \Omega$ for $-\varepsilon_0 < \varepsilon < \varepsilon_0$

it follows that

$$\Phi(\varepsilon) = g(x_0 + \varepsilon h + y(\varepsilon))$$

has an unconstrained extremum at 0 over the interval $(-\varepsilon_0, \varepsilon_0)$. By Theorem 1 (Lecture 4 Cont'd), it follows that

$$0 = \left.\frac{d}{d\varepsilon}\Phi(\varepsilon)\right|_{\varepsilon=0} = \left. Dg(x_0 + \varepsilon h + y(\varepsilon))\right|_{\varepsilon=0} \lim_{\varepsilon\to 0} \frac{\varepsilon h + y(\varepsilon)}{\varepsilon}$$

$$= Dg(x_0) h$$

$$\Rightarrow h \in \operatorname{Ker}(Dg(x_0)) \qquad \blacksquare$$